



TÉLUQ

Utiliser un logiciel antiplagiat : Pour ou contre ?

Serge Gérin-Lajoie
Professeur
Colloque REFAD 2018, Québec, 26 mai 2018



TÉLUQ


L'impact du numérique... en éducation

Les institutions d'enseignement sont actuellement confrontées à une réalité indéniable :

- Le développement du Web et de ses ressources a modifié radicalement la recherche documentaire et la réalisation des travaux académiques.
- Étudiants, professeurs et professionnels œuvrant au sein des institutions de formation recourent à Internet.

CREPUQ (2011)

Colloque REFAD 2018



L'impact du numérique... chez les étudiants


- 91 % des étudiants de 13-25 ans utilisent un ordinateur pour faire leurs travaux scolaires;
- 97 % des étudiants utilisent Internet comme source documentaire;
- 75 % ne donnent pas de références.

Nicole Perreault (2009)

- Tous les étudiants utilisent Internet pour se documenter quand seulement 1 étudiant sur 2 se rend encore en bibliothèque.

Six Degrés (2006)

Colloque REFAD 2018




L'impact du numérique... chez les étudiants

- L'accessibilité des informations fournies par Internet combinée à la facilité d'utiliser la fonction copier-coller aurait amplifié le phénomène du plagiat.
- Les étudiants pouvant accéder aisément aux données du Web, il leur semble « légitime » de se les approprier et de ne pas fournir les références nécessaires ni de demander les permissions d'utilisation.

CREPUQ (2011)

Colloque REFAD 2018




Le « copier-coller »

- Les fonctions **copier-coller** ont le mérite d'avoir affranchi tout rédacteur de la fastidieuse tâche de recopier mot à mot tout texte et qu'il souhaite emprunter.
- Ces fonctions s'inscrivent dans la suite logique d'avoir cherché et trouvé des informations ainsi que de les avoir jugées dignes de faire partie d'un travail à remettre.
- Il s'agit ici de compétences informationnelles aujourd'hui indispensables à tout professionnel, en devenir ou en emploi.

Morin (2014)

Colloque REFAD 2018



En tout cas, un phénomène en changement

Il y a 25 ans, le plagiat était le fait de deux étudiants qui copiaient leurs travaux et partageaient leurs notes d'année en année.

Aujourd'hui, le plagiat tend à l'utilisation de sites web qui proposent des travaux ou l'utilisation d'informations provenant de sites web comme Wikipedia. Les recherches d'informations spécifiques étant aidées par l'usage de moteurs de recherche comme Google.

Traduction libre de Briggs (2008)

Colloque REFAD 2018

TÉLUQ

Si les étudiants utilisent le numérique pour faire leurs travaux, pourquoi ne pas utiliser les mêmes outils pour détecter les cas de plagiat ?



Source: <http://biblio.bas-rhin.fr/actualites/inclusion-numerique-une-enquete-l-anah/>




Les logiciels de détection de plagiat

Comment ça marche ?

Habituellement, les logiciels de détection de plagiat fonctionnent ainsi:

1. Extrait le texte à partir des documents soumis (MS Word, PDF, etc.).
2. Recherche les similitudes entre ce texte et ce qui est disponible sur Internet ou dans les bases de données publiques ou privées.
3. Montre les résultats de l'analyse et quantifie la proportion du texte qui est similaire avec ce qui se retrouve dans les sources de comparaison.



Les logiciels de détection de plagiat

Comment ça marche ?

La base de comparaison

- Documents couramment indexés sur le Web;
- Travaux soumis par les étudiants dans les établissements qui utilisent le même logiciel;
- Articles de revue scientifique, des documents professionnels, mais qui ne sont pas relevés par les moteurs de recherche courants, notamment des contenus payants.


Audet (2011)

Colloque REFAD 2018





Source: <http://www.jeudubazar.com/en-famille/2567-nouvelles-histoires-extension-il-etait-une-fois-edition-2016-3558380032052.html>

Colloque REFAD 2018


TÉLUQ

Contexte

- Certains enseignants demandent de l'aide pour détecter le plagiat.
- Par le passé, l'établissement utilisait un système de détection de plagiat (Turnitin), mais plusieurs problèmes étaient rencontrés.
 - Anonymisation des copies (temps nécessaire).
 - Sécurité des serveurs (travaux sur serveurs aux États-Unis).



Source: <https://www.nawafedh.org/node/2432>

Colloque REFAD 2018
11


TÉLUQ

Contexte

Le mandat:

- Tester le logiciel (Ithenticate).
- Explorer ce qui existe comme alternatives gratuites.
- Proposer une offre de services qui pourrait être mise en place dans l'établissement.



Source: <https://www.nawafedh.org/node/2432>

Colloque REFAD 2018
12



Méthodologie

1. Ouverture d'un compte dans Ithenticate.
2. Achat de crédits.
3. Tester une thèse et un mémoire « louches » identifiés par des enseignants.
4. Comparaison des sections identifiées comme « plagiées » avec:
 - Google
 - Yahoo
 - Positéo (logiciel antiplagiat gratuit)



Source: <http://www.surveypolpole.fr/methodologie/>

13

Colloque REFAD 2018




Résultats



Match Overview

Match	Source	Percentage
1	Internet 612 words created on 13 Feb 2014 www.agripress.be	13%
2	Crossref 171 words Vanneste, Céline, Vervaeke, E. & Michels, E. (2014) 'Animal Manure: From Waste to Resource', <i>Renewable Energy</i> 68, 1-10	4%
3	Internet 130 words created on 27 Oct 2019 www.ars-antiparis.fr	3%
4	Crossref 107 words Rötter, R. & B. 'Thermochemical conversion of livestock waste: Carbonization of waste solids', <i>Bioresour. Technol.</i>	2%
5	Internet 62 words created on 16 May 2016 www.ars-antiparis.fr	1%
6	Internet 45 words created on 10 Mar 2012 www.ars-antiparis.fr	1%
7	Crossref 41 words Blomster, S. & G. 'Nutrient balance - From waste to Resource', <i>Renewable Energy</i> 68, 1-10	1%
8	Internet 36 words created on 25 Feb 2014 www.ars-antiparis.fr	1%
9	Internet 32 words created on 28 Mar 2014 www.ars-antiparis.fr	1%
10	Crossref 30 words Chou, H. & G. 'Nutrient balance - From waste to Resource', <i>Renewable Energy</i> 68, 1-10	1%
11	Internet 25 words created on 24 Feb 2016 www.ars-antiparis.fr	1%

3.1.2 Processing of animal manure surplus




TÉLUQ

Tests avec Ithenticate

Résultats

- Les résultats obtenus sont intéressants pour les enseignants.
 - Permettent d'identifier rapidement les éléments plagiés.
 - Permettent l'analyse des zones plagées en comparant avec les sources plagées.

Colloque REFAD 2018




TÉLUQ

Tests avec Ithenticate

Limites constatées

- Les % de plagiat sont relatifs, donc une analyse détaillée est nécessaire.
- Rapport difficilement exportable et peu utile.
- Analyse détaillée des éléments plagés prend du temps.
- Analyse détaillée difficile à faire pour les gens hors du domaine disciplinaire.
- Analyse détaillée des éléments plagés ne permet pas de sélectionner et d'exporter les zones plagées.

Colloque REFAD 2018




Tests avec Ithenticate

Limites constatées- suite

- Un compte permet d'accéder à tous les rapports d'analyse
 - Aspect de confidentialité difficile à respecter si plusieurs utilisateurs.
 - Difficile d'envisager un accès de multiples personnes (Directions, conseillers facultaires, etc.).

Colloque REFAD 2018



Tests avec Ithenticate

Limites constatées- suite

- Nécessite un temps de préparation des documents à analyser (anonymisation, nbr de mots, etc.).
- Coûts d'utilisation élevés:
 - 100\$ US par analyse.
 - Possibilité d'acheter 50 analyses pour 4000\$ US.



Colloque REFAD 2018

TÉLUQ

Tests avec Ithenticate

Résultats de la comparaison avec Google, Yahoo et Positéo

90-95 % des éléments plagiés ont été trouvés avec Google ou Positéo

Colloque REFAD 2018

TÉLUQ

Si on me demandait une recommandation aujourd'hui ... pour ceux qui envisagent d'implanter un outil de détection de plagiat.

**Que faire ?
Quoi recommander ?**



Colloque REFAD 2018

Source: <http://www.industrie-news.com/fin-des-tarifs-reglementes-que-faire-si-votre-copropriete-est-concernee.html>


TÉLUQ


Si on me demandait une recommandation aujourd'hui

...



Source: <https://www.youtube.com/watch?v=55f92s3KhHQ>
21

Colloque REFAD 2018



TÉLUQ

Les logiciels de détection de plagiat dans la littérature ?

Que dit la littérature scientifique:

- à propos des logiciels de détection de plagiat ?
- de leur efficacité à détecter le plagiat ?

Colloque REFAD 2018




Les logiciels de détection de plagiat dans la littérature ?

L'efficacité des outils de détection de plagiat peut être mesurée de deux manières:

1. Leur efficacité à identifier des extraits de textes « non originaux ».
2. Leurs effets sur le comportement des étudiants à produire des textes originaux et à citer correctement les travaux déjà existants.

Badge (2010)

Colloque REFAD 2018




Les logiciels de détection de plagiat dans la littérature ?

Les travaux de Debora Weber-Wulff (cité par Badge, 2010)

- Tests intensifs et répétés depuis 2004.
- Évalue:
 - « Utilisabilité » des logiciels.
 - Efficacité de la détection.
 - À partir de textes originaux, extraits de Wikipedia, textes traduits, textes en « courteline » sans citation, etc.

Colloque REFAD 2018




Les logiciels de détection de plagiat dans la littérature ?

Résultats de 2008

- Les deux logiciels les plus performants ne sont pas mentionnés en éducation.
 - Copyscape (70/80)
 - Fait pour vérifier l'originalité de pages web, mais peu pratique pour vérifier des travaux étudiants.
 - Plagiarism Detector (68/80)
 - Provenance douteuse du logiciel.
 - Danger de virus informatique.

Badge (2010)

Colloque REFAD 2018




Les logiciels de détection de plagiat dans la littérature ?

Résultats de 2008 - suite

- Les deux logiciels les plus utilisés dans le domaine de l'éducation n'obtiennent pas de "très bons" résultats
 - Safeassign (57/80).
 - Turnitin (45/80).

Badge (2010)

Colloque REFAD 2018



Les logiciels de détection de plagiat dans la littérature ?

Différents éléments considérés:
(CRÉPUQ, 2011; Duguest, 2008; Duteille, 2007; Duteille et Boulerie, 2012)

- Coûts;
- Fonctionnalités;
- Simplicité d'utilisation;
- Fiabilité de l'analyse;
- Clarté des résultats obtenus;
- Accès aux bases de données (publiques et privées);
- Disponibilités Windows ou MAC;
- Langue d'interface;
- Présence de publicités;
- Format de fichiers pris en charge;
- Sécurité des données.

Colloque REFAD 2018




Les logiciels de détection de plagiat dans la littérature ?

Différentes limites répertoriées
(CRÉPUQ, 2011; Duguest, 2008; Duteille, 2007; Duteille et Boulerie, 2012)

- Disponible qu'en Windows;
- Interface en anglais;
- Ne prend pas en compte les traductions;
- Ne reconnaît pas les tableaux / graphiques / images;
- Les documents nécessitant un accès payant ne sont pas accessibles (revues, banques de données, sites d'achats et d'échange de travaux);
- Archivage des travaux sur des serveurs aux États-Unis.

Colloque REFAD 2018



Les logiciels de détection de plagiat dans la littérature ?

Différentes autres limites répertoriées

- Les logiciels de détection prétendent détecter le plagiat, mais dans les faits, ils n'indiquent que la proportion de texte "non original".
Konstantinidis et al. (2013).
- Aucun d'entre eux ne peut indiquer si les étudiants ont été malhonnêtes dans la réalisation de leur travail.
Ledwith & Rísquez (2008), cités dans Konstantinidis et al (2013).

Colloque REFAD 2018



Les logiciels de détection de plagiat dans la littérature ?

Différentes autres limites répertoriées

- Une exploration des extraits des textes "plagiés" est nécessaire par les enseignants.
Anderson (2009), cité dans Konstantinidis et al. (2013).
- Un jugement final par les enseignants est nécessaire pour déterminer si les étudiants ont plagié ou non.
Badge, Cann, et Scott (2007), cités dans Konstantinidis et al. (2013).

Colloque REFAD 2018




Les logiciels de détection de plagiat dans la littérature ?

Différentes limites répertoriées

- Les recherches avec des outils comme Google sont plus intuitives pour identifier des cas de plagiat.
Evans (2006)
- Faire les analyses des travaux avec les logiciels antiplagiat augmente le travail des enseignants.
McKeever (2006)

Colloque REFAD 2018




Les logiciels de détection de plagiat dans la littérature ?

Résultats – recommandations

- Pour un usage ponctuel, les outils *Copyscape*, *Article Checker* et *Dupli Checker* peuvent convenir.
- Pour une solution plus fiable, polyvalente (types de format) et intuitive, mais à un coût non négligeable, retenir les solutions payantes : Ephorus, Turnitin, Compilatio et Urkund.
Duteille (2007) et Duteille & Boulerie (2012)

Colloque REFAD 2018




Les logiciels de détection de plagiat dans la littérature ?

L'efficacité des outils de détection de plagiat peut être mesurée de deux manières:

1. Leur efficacité à identifier des extraits de textes « non originaux ».
2. Leurs effets sur le comportement des étudiants à produire des textes originaux et à citer correctement les travaux déjà existants.

Badge (2010)

Colloque REFAD 2018




Les logiciels de détection de plagiat dans la littérature ?

Introduction de Turnitin à l'Université McGill en 2006-2007

- 25-30 enseignants
- Sondage auprès des étudiants
 - 34 % n'éviteraient pas de suivre un cours dans lequel un logiciel de détection des similitudes y est employé;
 - la qualité de la discussion au sujet des citations a été supérieure dans les cours où le logiciel est utilisé;
 - la question de la confiance entre étudiant et enseignant n'a pas été affectée;
 - qu'ils avaient pris plus de temps pour s'informer sur les attentes relatives à la manière de citer les sources que dans d'autres cours.

CREPUQ (2011)

Colloque REFAD 2018




Les logiciels de détection de plagiat dans la littérature ?

Introduction de Turnitin à l'Université McGill en 2006-2007

- Sondage auprès des enseignants:
 - Logiciel a fourni une rétroaction bénéfique aux étudiants;
 - Logiciel aide à comprendre la nécessité et la valeur de la citation;
 - Logiciel permet d'approfondir la discussion sur le plagiat et la copie.

CREPUQ (2011)

Colloque REFAD 2018



Les logiciels de détection de plagiat dans la littérature ?

L'introduction de logiciels antiplagiat aide à réduire le plagiat chez les étudiants.

Bilic-Zulle, et al. (2008), Ledwith & Rasquez, 2008, cités dans Konstantinidis et al. (2013)

PR Newswire Association (2006) reports that "institutions adopting Turnitin observed an immediate deterrent effect, resulting in levels of plagiarized student work dropping to approximately 30% from 77%".

Köse & Arikan (2011).

Colloque REFAD 2018

TÉLUQ

Les logiciels de détection de plagiat dans la littérature ?

En bref

« L'efficacité technique des logiciels antiplagiat est peut-être faible, mais leur efficacité objective est apparemment indéniable, comme celle d'un vieux gardien à l'entrée ».
Lamontagne (2011)

« Ils ont le mérite de mettre en garde les étudiants ».
Bergadaà (2015)

Colloque REFAD 2018

TÉLUQ

Pour ceux qui envisagent d'implanter un outil de détection de plagiat.



Source: <http://patrimoinepq.blogspot.ca/2010/07/finalement-le-film-selections.html>

Colloque REFAD 2018

38




Les logiciels de détection de plagiat

Recommandations

1. Proposer aux enseignants de travailler avec Google ou Positéo.
2. Travailler avec les facultés pour développer des banques de travaux à partir des boîtes de dépôt du Portail des cours.
 - Comparaisons internes des travaux avec des logiciels comme : WCopyfind (PC) ou Novus Scan (Mac).
 - Éventuellement, utiliser un logiciel antiplagiat avec une telle banque.

Colloque REFAD 2018




Les logiciels de détection de plagiat

Recommandations - suite

3. Proposer l'utilisation du logiciel Ithenticate ou Plagscan à la pièce.
4. Limiter les travaux analysés aux mémoires et thèses.
5. Fournir un accès au logiciel pour le commissaire aux infractions afin de consulter les rapports afin de mener ses enquêtes.

Colloque REFAD 2018



Les solutions

- Mise en place d'une approche intégrée, dans laquelle on retrouve l'utilisation d'un logiciel de détection de similitudes;
- Sensibilisation (discussion en salle de classe, tests, etc.);
- Mise en place de cours ou d'ateliers de méthodologie obligatoires;
- Vérifications des travaux par étudiants;
- Suivi auprès des plagiaires;
- Exigences pour les travaux rendant le plagiat difficile;
- Etc.

Colloque REFAD 2018




Merci beaucoup

Je suis ouvert à vos questions et commentaires





Professeur
serge.gerin-lajoie@teluq.ca



Références

Audet, L. (2011). *Les pratiques et défis de l'évaluation en ligne*. Disponible à : http://archives.refad.ca/evaluation_en_ligne.pdf

Badge, J. (2010). How effective are electronic plagiarism detection systems and does it matter how you use them? Reviewing the evidence. In *Plagiarism conference 2010*.


Bergadaà, M. (2015). Une brève histoire de la lutte contre le plagiat dans le monde académique. *Questions de Communication*, 27(1), 171–188.

Briggs, L. (2008). Thwarting the copycats. *Campus Technology*, 22(1), 22–28. Disponible à : <https://campustechnology.com/Articles/2008/09/Thwarting-the-Copycats.aspx>

Commission de l'éthique de la science et de la technologie. (2005). *Le plagiat électronique dans les travaux scolaires : Pour une réflexion éthique*.

Conférence des recteurs et principaux des universités du Québec. (2011). *Les logiciels de détection de similitudes : une solution au plagiat électronique ?*

Colloque REFAD 2018



Références

Duguest, D. (2008). *Etude comparative des logiciels anti plagiat*. Nantes.

Duteille, O. (2007). *Comparatif Logiciels anti-plagiat*.

Duteille, O., & Boulerie, P. (2012). *Comparatif Logiciels anti-plagiat*.


Evans, R. (2006). Evaluating an electronic plagiarism detection service. *Active Learning in Higher Education*, 7(1), 87–99. <https://doi.org/10.1177/1469787406061150>

Heather, J. (2010). Turnitoff: identifying and fixing a hole in current plagiarism detection software. *Assessment & Evaluation in Higher Education*, 35(6), 647–660. <https://doi.org/10.1080/02602938.2010.486471>

Konstantinidis, A., Theodosiadou, D., & Pappos, C. (2013). Plagiarism: Examination of Conceptual Issues and Evaluation of Research Findings on Using Detection Services. *Contemporary Educational Technology*, 4(3), 212–221.

Köse, Ö., & Arikan, A. (2011). Reducing plagiarism by using online software: An experimental study. *Contemporary Oline Language Education Journal*, (1), 122–129.

Colloque REFAD 2018



Références

Lamontagne, D. (2011). L'efficacité des logiciels anti-plagiat. In C. Vaufrey (Ed.), *Respecter le droit d'auteur, prévenir le plagiat*. Thot Cursus. Disponible à : <http://cursus.edu/media/upload/livre-blanc-plagiat-final2.pdf>.

McKeever, L. (2006). Online plagiarism detection services—saviour or scourge? *Assessment & Evaluation in Higher Education*, 31(2), 155–165.

Morin, S. (2014). « Copier-coller-citer »! *L'éveilleur*. Disponible à : <https://leveilleur.espaceweb.usherbrooke.ca/19550/copier-coller-citer/>

Perreault, N. (2009). Portrait et enjeux du plagiat électronique dans les universités québécoises. In *Le plagiat dans les universités québécoises à l'ère du numérique*. Retrieved from http://www.powershow.com/view1/298e32-ZDc1Z/Portrait_et_enjeux_du_plagiat_lectronique_dans_les_universits_qubcoises_p_owerpoint_ppt_presentation

Six degrés. (2006). *Les usages d'Internet dans l'enseignement supérieur : « De la documentation... au plagiat »*.

Colloque REFAD 2018



TÉLUQ

teluq.ca